Understanding Risk in a Biopharmaceutical Portfolio

By

Alice Elizabeth Wagner

MA, Clinical Investigation, Boston University, 2006
BS, Pre-professional Studies, University of Notre Dame, 2002

SUBMITTED TO THE HARVARD – MIT DIVISION OF HEALTH SCIENCES AND TECHNOLOGY AND THE MIT SLOAN SCHOOL OF MANAGEMENT IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREES OF

MASTERS OF SCIENCE IN HEALTH SCIENCES AND TECHNOLOGY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2011 © 2011 Joe B. Gillis. All rights reserved.

ARCHIVES

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

SEP 2 1 2011

LIBRARIES

The author hereby grants MIT permission to reproduce and distribute publicly paper and electronic copies of this thesis document in whole or in part.

Signature of Author: _____ Harvard–MIT Division of Health Sciences and Technology and MIT Sloan School of Management, August 2011

Certified by: _____ Fiona Murray, PhD
Associate Professor of Technological Innovation, Entrepreneurship, and Strategic Management, MIT Sloan School of Management

Certified by: _____ Andrew W. Lo
Harris & Harris Group Professor of Finance, Director of the Laboratory for Financial Engineering, MIT Sloan School of Management

Accepted by: _____ Ram Sasisekharan, PhD
Director, Harvard-MIT Division of Health Sciences and Technology
Edward Hood Taplin Professor of Health Sciences & Technology and Biological Engineering

MIT Libraries
Document Services

Room 14-0551
77 Massachusetts Avenue
Cambridge, MA 02139
Ph: 617.253.2800
Email: docs@mit.edu
http://libraries.mit.edu/docs

DISCLAIMER OF QUALITY

Due to the condition of the original material, there are unavoidable flaws in this reproduction. We have made every effort possible to provide you with the best copy available. If you are dissatisfied with this product and find it unusable, please contact Document Services as soon as possible.

Thank you.

Pages 65-70 contain illegible text. This is the best copy available.

Understanding Risk in a Biopharmaceutical Portfolio

by

Alice Elizabeth Wagner

Submitted to the Harvard-MIT Division of Health Sciences and Technology
on September 6, 2011, in partial fulfillment of the
requirements of the degree Masters of Science

Abstract

Investors have difficulty funding the life sciences because of the high risks involved in research and development and commercialization of new products. Risk in the biopharmaceutical industry is the result of scientific, regulatory and economic uncertainty. The nature of the biopharmaceutical industry introduces many challenges. Each of these challenges incorporates a measure of risk into drug development. The level of understanding of technical success interdependencies has not been fully investigated. These interdependencies (correlations) could lead to an overall greater risk to the company's portfolio than previously expected. A better understanding of the risks that lead to success or failure in drug development might encourage more investment in the life sciences and specifically in the biopharmaceutical industry, and a greater awareness of the correlations between risks and products might lead to more informed decision making on a biopharmaceutical portfolio leading increased productivity. A dataset was collected from Thomson Reuters. The dataset is the oncology portfolio from a biopharmaceutical company, Genentech Inc. Logistic regression was used to determine if any of the defined variables contributed to the success or failure of the oncology products. The chi-square value was 7.738 with the degrees of freedom equal to 5 and with a p-value of 0.17. Therefore, none of the variables significantly contributed to the outcome. More research should be performed in this area in order to better understand the risk in a biopharmaceutical portfolio.

Thesis Advisor: Fiona Murray, PhD
Title: Associate Professor of Technological Innovation,
Entrepreneurship, and Strategic Management

Thesis Advisor: Andrew W. Lo
Title: Harris & Harris Group Professor of Finance,
Director of the Laboratory for Financial Engineering

Acknowledgements

I would like to thank the following people for supporting the completion of this thesis. Without them, this work would never have happened. Thank you.

Professor Lo and Professor Murray for advising me.

The people who invest in healthcare and who work in the healthcare field for devoting their lives to the betterment of mankind.

My friends in BEP – Jonathan Alspaugh, Alan Braly, Josh Gottlieb and Julie Hermann – for brainstorming ideas and for sharing in the frustrations and joys over the last three years.

David Pelly for his different way of thinking about the analytics.

My family for supporting every decision I have ever made and for providing me with unconditional love.

Table of Contents

List of Figures

List of Tables

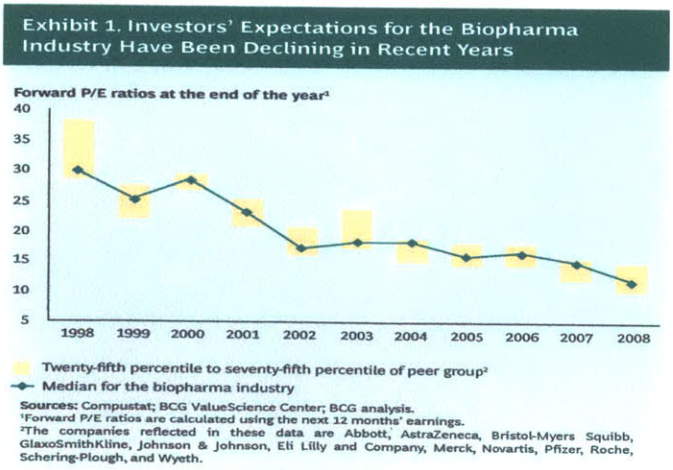
Chapter 1: Investing in the Life Sciences

History of Life Science Investing

The biopharmaceutical ecosystem is comprised of a complex web of interdependent organizations. These include drug development organizations such as universities, the National Institute of Health (NIH), Research and Development (R&D boutiques, established biopharmaceutical companies, and health care providers, and investment organizations such as governments, non-profits, angel funds, venture capital firms, private equity firms and pension funds.

Investors have reevaluated their beliefs about the efficiency of biopharmaceutical's innovation engine – and hence about the entire future value of the industry. Figure 1 shows the steady decline of the industry's price-to-earnings (P/E) ratio, a measure of investor confidence in future profit growth.[1]

Figure 1: Investors' Expectations for the Biopharma Industry



Public Funding in the Life Sciences

The major source of public funding in the life sciences in the United States (US) is from the National Institute of Health (NIH). The NIH invests over $31.2 billion annually in medical research. More than 80% of the NIH's funding is awarded through almost 50,000 competitive grants to more than 325,000 researchers at over 3,000 universities, medical schools, and other research institutions in every state and around the world. About 10% of the NIH's budget supports projects conducted by nearly 6,000 scientists in its own laboratories, most of which are on the NIH campus in Bethesda, Maryland[2]. Table 1 shows the allocation of NIH extramural awards from 2001-2010.

Table 1: NIH Extramural Awards[a]

Number of Awards and Organizations Funded by Organization Type

Fiscal Years* 2001-2010

Fiscal Year	Organization Type	Number of Awarded Organizations	Number of Awards[b]	Total Funding[c]
2001	Higher Education	516	36,277	$12,278,336,624
2001	Research Institutes	200	3,437	$1,641,875,390
2001	Independent Hospitals	140	3,585	$1,301,412,303
2001	Other Domestic Nonprofit	305	984	$398,819,432
2001	Domestic For-Profit	1,287	2,180	$1,082,546,743
2001	Foreign (includes all institution types listed above)	165	368	$77,864,500

2001	TOTAL	2,613	46,831	$16,780,854,992
2002	Higher Education	515	38,361	$13,904,024,894
2002	Research Institutes	205	3,669	$1,825,220,560
2002	Independent Hospitals	140	3,774	$1,458,491,086
2002	Other Domestic Nonprofit	312	1,068	$490,742,491
2002	Domestic For-Profit	1,420	2,373	$1,276,015,658
2002	Foreign (includes all institution types listed above)	223	471	$119,970,107
2002	**TOTAL**	**2,815**	**49,716**	**$19,074,464,796**
2003	Higher Education	534	40,516	$15,665,006,383
2003	Research Institutes	208	3,889	$2,085,466,889
2003	Independent Hospitals	138	3,750	$1,547,434,901
2003	Other Domestic Nonprofit	330	1,184	$556,865,509
2003	Domestic For-Profit	1,600	2,700	$1,814,251,570
2003	Foreign (includes all institution types listed above)	273	548	$197,773,159
2003	**TOTAL**	**3,083**	**52,587**	**$21,866,798,411**
2004	Higher Education	532	41,554	$16,102,526,149
2004	Research Institutes	210	3,897	$2,183,670,516
2004	Independent Hospitals	124	3,842	$1,611,238,994
2004	Other Domestic Nonprofit	338	1,245	$595,596,475
2004	Domestic For-Profit	1,684	2,880	$2,054,962,258
2004	Foreign (includes all institution types listed above)	294	603	$352,582,195
2004	**TOTAL**	**3,182**	**54,021**	**$22,900,576,587**
2005	Higher Education	535	41,570	$16,688,059,009
2005	Research Institutes	211	3,924	$2,158,045,271
2005	Independent Hospitals	117	3,993	$1,708,661,086

2005	Other Domestic Nonprofit	373	1,263	$546,051,495
2005	Domestic For-Profit	1,851	3,022	$2,025,019,484
2005	Foreign (includes all institution types listed above)	333	793	$284,337,954
2005	**TOTAL**	**3,419**	**54,564**	**$23,410,118,044**
2006	Higher Education	542	41,479	$16,367,084,439
2006	Research Institutes	207	3,914	$2,193,172,719
2006	Independent Hospitals	116	4,014	$1,731,033,176
2006	Other Domestic Nonprofit	379	1,222	$532,030,211
2006	Domestic For-Profit	1,878	4,060	$2,066,408,275
2006	Foreign (includes all institution types listed above)	327	744	$293,231,098
2006	**TOTAL**	**3,449**	**55,433**	**$23,182,959,918**
2007	Higher Education	532	41,767	$16,801,934,311
2007	Research Institutes	200	3,812	$2,104,988,996
2007	Independent Hospitals	108	3,936	$1,763,529,298
2007	Other Domestic Nonprofit	346	1,185	$615,363,140
2007	Domestic For-Profit	1,634	2,688	$2,008,682,906
2007	Foreign (includes all institution types listed above)	315	643	$206,892,687
2007	**TOTAL**	**3,125**	**54,031**	**$23,501,391,338**
2008	Higher Education	524	41,110	$16,779,594,423
2008	Research Institutes	203	3,656	$2,134,144,227
2008	Independent Hospitals	101	3,899	$1,731,243,120
2008	Other Domestic Nonprofit	325	1,096	$481,153,827
2008	Domestic For-Profit	1,385	2,310	$2,088,855,788
2008	Foreign (includes all institution types listed above)	309	666	$277,239,785

2008	TOTAL	2,846	52,737	$23,492,231,170
2009	Higher Education	506	40,379	$17,285,925,798
2009	Research Institutes	196	3,458	$2,185,767,859
2009	Independent Hospitals	96	3,744	$1,727,540,656
2009	Other Domestic Nonprofit	308	1,084	$605,253,735
2009	Domestic For-Profit	1,374	2,264	$2,390,424,600
2009	Foreign (includes all institution types listed above)	288	616	$282,331,162
2009	**TOTAL**	**2,768**	**51,545**	**$24,477,243,810**
2010	Higher Education	478	39,951	$16,953,363,969
2010	Research Institutes	185	3,327	$1,941,045,871
2010	Independent Hospitals	89	3,760	$1,751,914,669
2010	Other Domestic Nonprofit	251	1,017	$522,949,207
2010	Domestic For-Profit	1,247	1,867	$853,818,657
2010	Foreign (includes all institution types listed above)	241	516	$211,465,105
2010	**TOTAL**	**2,491**	**50,438**	**$22,234,557,478**

[a] **NIH Awards** — Includes all grants and contracts except 2010 for which contract data is not included.

[b] **Number of Awards** — The number of awards is intended to show the number of unique projects funded. Therefore, for grants, the number of noncompeting supplements is not included in the number of awards because these supplements support existing projects, without expanding the scope of work. However, the award amounts for noncompeting supplements are included in the award amount, because they reflect total expenditures on funded projects. In contrast, the number of competing supplements is included in the number of awards because these supplements represent expanded scope of work on existing projects; the award amounts for competing supplements are also included in the award amount. Similarly, for contracts, the number of noncompeting modifications is not included in the number of awards because these modifications support existing projects, without expanding the scope of work. The award amounts for noncompeting modifications are included in the award amount, because they reflect total expenditures on funded projects.

[c] **Total Funding** — Total funding is the funding amount for each fiscal year, and not for the life of the project. Includes only awards made with Direct Budget Authority, Superfund Budget Authority and Reimbursable funds.

*Due to the use of more refined analysis techniques, some pre-2009 data previously published has been updated in this table.

Data drawn from frozen FY 2010 Pub file as of 12/08/2010.

Private Funding in the Life Sciences

Investing in the life sciences through private methods typically involves one or a combination of the following vehicles: angel funds, venture capital funds, private equity funds, pension funds, or established biopharmaceutical companies.

There are a number of funding stages that typically describe the maturity of a company. Seed stage financing is a small initial investment, usually under $1 million, used to validate a concept, get a company formed and complete the initial business plan. Series A/B financing is one or two early rounds, roughly $1-5 million for Series A and $6-10 million for Series B, that are typically venture capital financed. Series C/D financing are possible financing rounds, generally $15-50 million, intended to take a company through an exit. Mezzanine financing is classically the last financing round and the size depends upon the needs of the company before the exit. A company might require bridge financing before the completion of another round of financing or before an IPO to tie them over. Finally, a buyout occurs when a company is purchased by venture capital firm or investor group, after which the incumbent and/or incoming management will be given or acquire a large stake in the business[3].

Angel Investors

An angel investor or angel (also known as a business angel or informal investor) is an affluent individual who provides capital for a business start-up, usually in exchange for convertible debt or ownership equity. A small but increasing number of angel investors organize themselves into angel groups or angel networks to share research and pool their

investment capital[4].

Angel capital fills the gap in start-up financing between "friends and family" who provide seed funding, and venture capital. Angel investments bear extremely high risk and are usually subject to dilution from future investment rounds. As such, they require a very high return on investment[5]. Since a large percentage of angel investments are lost completely when early stage companies fail, professional angel investors seek investments that have the potential to return at least 10 or more times their original investment within 5 years, through a defined exit strategy. According to the Center for Venture Research, there were 258,000 active angel investors in the US in 2007[6]. The past few years, particularly in North America, have seen the emergence of networks of angel groups, through which companies that apply for funding to one group are then brought before other groups to raise additional capital[7].

Venture Capital

Money provided by investors to startup firms and small businesses with perceived long-term growth potential. This is a very important source of funding for startups that do not have access to capital markets. It typically entails high risk for the investor, but it has the potential for above-average returns[8].

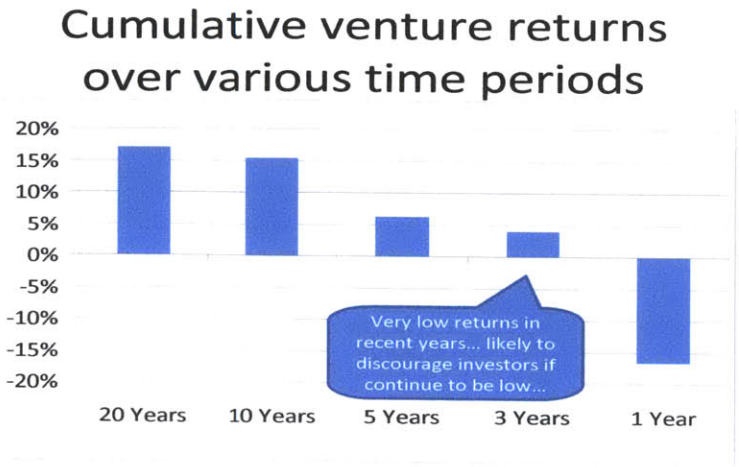
Venture capitalists are often wary of investing in biotechnology because bioentrepreneurs seldom provide realistic estimates of the value of their technologies[9]. Investing in healthcare has a higher bar than investing in other industries. Mike Carusi of Advanced

Technology Ventures states, "A lot of biotech firms can't get there. First off, they take too long. Second, you have to invest too much money before you know you actually have something. And finally, the technology risk is too high. Only one out of 10 drugs makes it through to final approval."[10]

Not all VC firms even invest in health care. Investors do not get the same absolute return in dollars on the health care that you do on telecoms or the Internet. "In venture capital, the health care only firms are the ones that might invest in early-stage biotech (therapeutics, i.e. drug development) companies, while a balanced fund like ourselves won't," says Mike Carusi.[10]

Pension funds contribute to over 50% of US venture funds[11]. Outside investors will put money into venture funds only if they expect returns (IRR) to be greater than the cost of capital. According to a study performed by Cambridge Associates covering 1606 biotechnology companies from 1986-2008, venture investment resulted in an average gross IRR on realized biopharmaceutical investments of 20.7%, after fees and other costs (5%). However, this came at huge risks. Of those investments, 44% were a full or partial loss (Figure 2), two-thirds of the profitable investments took 5 years or more to be realized, and over 1200 data points had yet to pay out by December 31, 2008[12].

Figure 2: Cambridge Associates Study: Cumulative Venture Returns



Source: Thomson VentureXpert. All returns for the period ending December 31, 2008. Net return to investors in VC funds.

Risks in Investing in the Life Sciences

Investors have difficulty funding the life sciences because of the high risks involved in research and development and commercialization of new products. These risks include high cost of capital, technical uncertainty (fewer than 1% of drug candidates will make it to market), regulatory uncertainty, long time lines (typically 10 years or more) and many others.

Risk in the biopharmaceutical industry is the result of scientific, regulatory and economic uncertainty. The first two risks create the lengthy development time and thereby the economic risk. The longer the scientific development time, the greater the likelihood that a competitor will make the discovery first and thereby greatly diminish the possibility for a return on the R&D investment of the innovator. Regulatory uncertainty occurs because

the time required for new drug approval further delays product marketing, and because marketing approval is not assured[13].

Cost of capital

Investments in R&D are delivering diminished returns. A mere five years ago, industry executives could expect every dollar invested in discovering new therapeutics to yield a risk-adjusted return of 15% or more. Yet today, despite a wealth of new scientific discoveries, returns on R&D in biopharmaceuticals have fallen to 11% or less, a rate barely covering the cost of capital[1].

Evidence shows that the cost of capital for venture backed early stage companies in life sciences is high (many estimates suggest 20% or higher). This reflects investors' expectation of a return sufficient to compensate them for taking on extraordinary risk. The cost of capital is a critical benchmark for assessing commercial viability of a project; it measures the opportunity cost of resources, it is often used as the hurdle rate of return to decide whether to invest, and it is also used as a discount rate to evaluate future cash flows. Outside investors will put money into venture funds only if they expect returns (IRR) to be greater than the cost of capital[1].

Technical Uncertainty

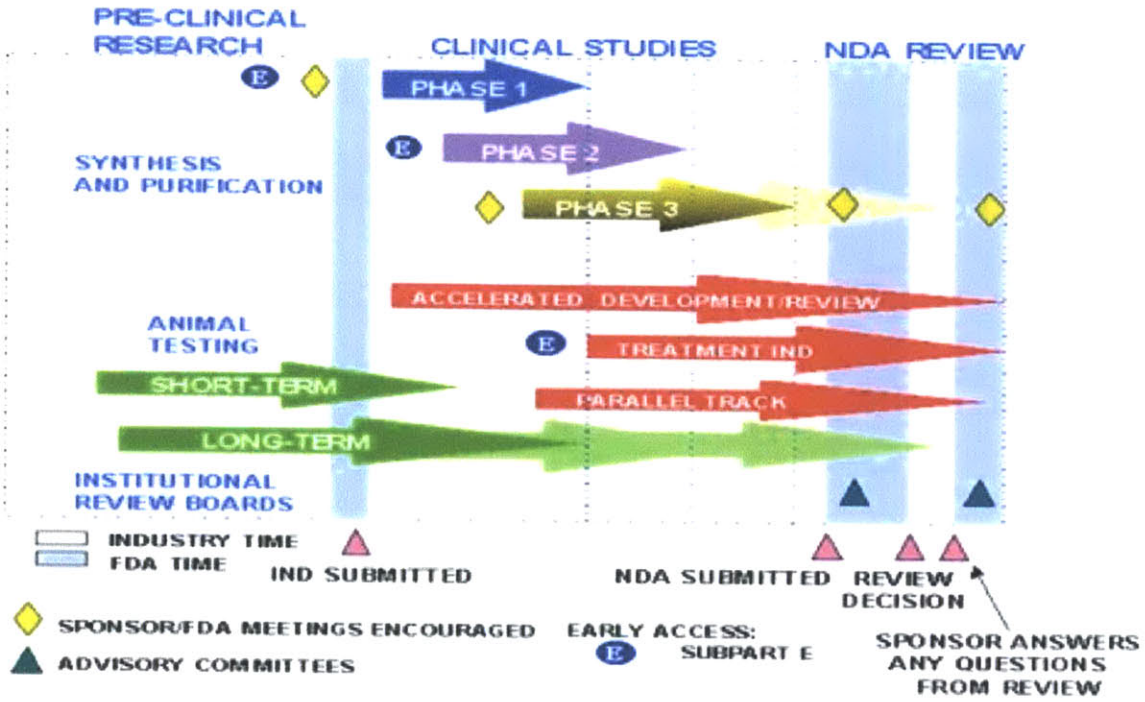
The concept of technical uncertainty is borrowed from Robert Pindyck. Technical uncertainty arises from the physical difficulty of completing an R&D project. At the origination of the project, limited information is available regarding the efforts, resources and time required to successfully realize the future return. Initiating the project and completing successive stages will incrementally reveal information related to these issues. As the investment proceeds, the barriers to completion may become higher or lower, but the true cost of the investment is only known with certainty once the projected has reached completion. Thus, the "information revealing" nature of the technical uncertainty enhances the incentives to commence investment[14].

There are several components to technical risk. A breakdown might look like:

- Known technical risks: safety (toxicity), efficacy, basic mechanism of action (same target), clinical trial endpoint selection
- Unknown natural (potentially correlated) risks: dose-response relationship, biomarkers for target, biomarkers for dose, established standard of care

Each stage of drug development has its own risks that role up into the total risk profile for a given R&D project. The phases of drug development are portrayed in Figure 3.

Figure 3: Stages of Drug Development[15]



Below is a description of each stage of drug development.

Preclinical stage:

This stage comprises of study on animals to find out various parameters for a drug under development. During preclinical drug development, a sponsor evaluates the drug's toxic and pharmacological effects through in vitro and in vivo laboratory animal testing. Genotoxicity screening is performed, as well as investigations on drug absorption and metabolism, the toxicity of the drug's metabolites, and the speed with which the drug and its metabolites are excreted from the body. At the preclinical stage, the Food & Drug

Administration (FDA) will generally ask, at a minimum, that sponsors: (1) develop a pharmacological profile of the drug; (2) determine the acute toxicity of the drug in at least two species of animals, and (3) conduct short-term toxicity studies ranging from 2 weeks to 3 months, depending on the proposed duration of use of the substance in the proposed clinical studies[16].

Clinical Stages:

Phase I: Phase I studies are carried out in healthy volunteers, which are small in number – usually 20 to 100. The purpose of phase I studies is to identify metabolic and pharmacological effects of drug in humans and to determine the side effects associated with increasing doses, and, if possible, to gain early evidence on effectiveness. During Phase 1, sufficient information about the drug's pharmacokinetics and pharmacological effects is required. The purpose of phase I studies is to mainly determine safety profile[16].

Phase II: Phase 2 includes the early controlled clinical studies conducted to obtain some preliminary data on the effectiveness of the drug for a particular indication or indications in patients with the disease or condition. This phase of testing also helps determine the common short-term side effects and risks associated with the drug. Phase 2 studies are typically well-controlled, closely monitored, and conducted in a relatively small number of patients, usually involving several hundred people[16].

Phase III: Phase 3 studies are expanded controlled and uncontrolled trials. They are performed after preliminary evidence suggesting effectiveness of the drug has been obtained in Phase 2, and are intended to gather the additional information about effectiveness and safety that is needed to evaluate the overall benefit-risk relationship of the drug. Phase 3 studies also provide an adequate basis for extrapolating the results to the general population and transmitting that information in the physician labeling. Phase 3 studies usually include several hundred to several thousand people [16].

Phase IV: In addition to these three phases, Phase IV, also known as Post Marketing Surveillance is also carried out once the drug is approved and marketed. The aim of Phase IV is to find out safety profile in large patient pool across the world and to establish the safety profile of the drug. It is estimated that success rate of drugs making to market from lab is very less. One drug, from among the thousands tested, makes it to the market[16].

One can think of the technical risk as a series of wagers. Clinical drug development should be regarded as a series of high-risk wagers where success in the first wager (e.g. a phase I trial) allows a company to make additional wagers (e.g. phase II and phase III trials) before reaching the ultimate payoff (e.g. a marketed drug)[9].

The largest share of R&D spend by The Pharmaceutical Research and Manufacturers of America (PhRMA) members is in the Phase III and prehuman/preclinical trials. Over 20% of drug failures in drug development have been attributed to toxicity concerns. The

application of genomic technologies, such as biomarkers, offers the opportunity to save up to 34% of R&D costs[17].

Regulatory Uncertainty

The concept of regulatory uncertainty is also borrowed from Pindyck. Regulatory uncertainty arises from unpredictable aspects of the regulatory regime governing the completion of the R&D program, which may take the form of unpredictable compliance costs incurred over the course of the R&D project. The level of these costs may be higher or lower depending on how regulators respond to factors such as public opinion or safety concerns. In contrast to technical uncertainty, information about the regulatory regime proceeds regardless of whether or not the firm is investing. This tends to have a dampening effect on investment incentives, since the firm may benefit from delaying investments in order to observe the regulatory regime and thus obtain more information about its future trajectory[14].

The FDA has strict regulations involving laboratories, manufacturing, and clinical trials. The Code of Federal Regulations provides a detailed list of all of the rules that must be followed when developing a drug, biologic or device[15]. Adhering to these regulations involves substantial costs that must be factored into the investment decision.

Long timelines

Development times have increased over the last decade, which is largely due to drugs becoming more innovative and subsequently requiring more testing and developing compared to less novel drugs with familiar mechanisms of action[17].

Other risks

Getting a drug on the market is no longer just a matter of proving its clinical efficacy – a company must now prove the product's cost effectiveness as well[1]. Decisions about which health-care interventions represent adequate value to collectively funded health-care systems are as widespread as they are unavoidable. In the case of new biopharmaceuticals, many countries now require formal cost-effectiveness analysis to inform this decision-making process. This requires evidence on parameters associated with health-related utilities, treatment effects, resource use, and costs, for which data from available regulatory trials are invariably absent or highly uncertain. This uncertainty results from a number of factors including the predominance of intermediate end points in the clinical evidence-base and the limited period of follow-up of patients in clinical studies.

Despite these imperfections in the evidence base, decisions about whether new pharmaceuticals are sufficiently cost-effective for reimbursement cannot be side-stepped. Data limitations do, however, require the use of rigorous analytical methods to support decision making. Probabilistic decision models and value of information analysis offer a means of structuring decision problems, synthesizing all available data, characterizing the

uncertainty in the decision, quantifying the cost of uncertainty, and establishing the expected value of perfect information.

This analytical framework is important because it addresses two fundamental questions about new pharmaceuticals. First, is the product expected to be cost-effective on the basis of existing evidence? Second, is additional research concerning the product itself cost-effective? In addressing these questions, the analytical framework can establish when sufficient evidence exists to sustain a claim for a new pharmaceutical to be cost-effective[18].

Chapter 2: Biologic Basis of Disease - Oncology

Oncology Overview

Cancer is an abnormal growth of cells. Cancer cells rapidly reproduce despite restriction of space, nutrients shared by other cells, or signals sent from the body to stop reproduction. Cancer cells are often shaped differently from healthy cells, they do not function properly, and they can spread to many areas of the body. Tumors, abnormal growth of tissue, are clusters of cells that are capable of growing and dividing uncontrollably; their growth is not regulated.

Oncology is the study of cancer and tumors. The term "cancer" is used when a tumor is malignant, which is to say it has the potential to cause harm, including death. Tumors can be benign (noncancerous) or malignant (cancerous). Benign tumors tend to grow slowly and do not spread. Malignant tumors can grow rapidly, invade and destroy nearby normal tissues, and spread throughout the body. The original tumor is called the "primary tumor." Its cells, which travel through the body, can begin the formation of new tumors in other organs. These new tumors are referred to as "secondary tumors."

Cancer is not just one disease but rather a group of diseases, all of which cause cells in the body to change and grow out of control. Cancers are classified either according to the kind of fluid or tissue from which they originate, or according to the location in the body where they first developed. In addition, some cancers are of mixed types.

The following five broad categories indicate the tissue and blood classifications of cancer:

Carcinoma: A carcinoma is a cancer found in body tissue known as epithelial tissue that covers or lines surfaces of organs, glands, or body structures. For example, a cancer of the lining of the stomach is called a carcinoma. Many carcinomas affect organs or glands that are involved with secretion, such as breasts that produce milk. Carcinomas account for 80 percent to 90 percent of all cancer cases.

Sarcoma: A sarcoma is a malignant tumor growing from connective tissues, such as cartilage, fat, muscle, tendons, and bones. The most common sarcoma, a tumor on the bone, usually occurs in young adults. Examples of sarcoma include osteosarcoma (bone) and chondrosarcoma (cartilage).

Lymphoma: Lymphoma refers to a cancer that originates in the nodes or glands of the lymphatic system, whose job it is to produce white blood cells and clean body fluids, or in organs such as the brain and breast. Lymphomas are classified into two categories: Hodgkin's lymphoma and non-Hodgkin's lymphoma.

Leukemia: Leukemia, also known as blood cancer, is a cancer of the bone marrow that keeps the marrow from producing normal red and white blood cells and platelets. White blood cells are needed to resist infection. Red blood cells are needed to prevent anemia.

Platelets keep the body from easily bruising and bleeding. Examples of leukemia include acute myelogenous leukemia, chronic myelogenous leukemia, acute lymphocytic leukemia, and chronic lymphocytic leukemia. The terms myelogenous and lymphocytic indicate the type of cells that are involved.

Myeloma: Myeloma grows in the plasma cells of bone marrow. In some cases, the myeloma cells collect in one bone and form a single tumor, called a plasmacytoma. However, in other cases, the myeloma cells collect in many bones, forming many bone tumors. This is called multiple myeloma.

There is no one single cause for cancer. Scientists believe that it is the interaction of many factors together that produces cancer. The factors involved may be genetic, environmental, or constitutional characteristics of the individual.

Diagnosis, treatment, and prognosis for childhood cancers are different than for adult cancers. The main differences are the survival rate and the cause of the cancer. The survival rate for childhood cancer is about 75 percent, while in adult cancers the survival rate is 60 percent. This difference is thought to be because childhood cancer is more responsive to therapy, and a child can tolerate more aggressive therapy.

Childhood cancers often occur or begin in the stem cells, which are simple cells capable

of producing other types of specialized cells that the body needs. A sporadic (occurs by chance) cell change or mutation is usually what causes childhood cancer.

In adults, the type of cell that becomes cancerous is usually an "epithelial" cell, which is one of the cells that line the body cavity, including the surfaces of organs, glands, or body structures, and cover the body surface. Cancer in adults usually occurs from environmental exposures to these cells over time. Adult cancers are sometimes referred to as "acquired" for this reason.

The discovery of certain types of genes that contribute to cancer has been an extremely important development for cancer research. Over 90 percent of cancers are observed to have some type of genetic alteration. A small percentage (5 percent to 10 percent) of these alterations are inherited, while the rest are sporadic, which means they occur by chance or occur from environmental exposures (usually over many years).

There are three main types of genes that can affect cell growth, and are altered (mutated) in certain types of cancers, including the following:

Oncogenes: These genes regulate the normal growth of cells. Scientists commonly describe oncogenes as similar to a cancer "switch" that most people have in their bodies. What "flips the switch" to make these oncogenes suddenly become unable to control the normal growth of cells and allowing abnormal cancer cells to begin to grow, is unknown.

Tumor suppressor genes: These genes are able to recognize abnormal growth and reproduction of damaged cells, or cancer cells, and can interrupt their reproduction until the defect is corrected. If the tumor suppressor genes are mutated, however, and they do not function properly, tumor growth may occur.

Mismatch-repair genes: These genes help recognize errors when DNA is copied to make a new cell. If the DNA does not "match" perfectly, these genes repair the mismatch and correct the error. If these genes are not working properly, however, errors in DNA can be transmitted to new cells, causing them to be damaged.

Usually the number of cells in any of our body tissues is tightly controlled so that new cells are made for normal growth and development, as well as to replace dying cells. Ultimately, cancer is a loss of this balance due to genetic alterations that "tip the balance" in favor of excessive cell growth.

Pathways: Mechanisms of Action

There are many different ways in which a cancer can form. Ten signaling pathways have been identified, and multiple proteins are associated with each pathway. These proteins are either oncogenes or tumor suppressor genes. Characterization of these pathways has led scientists to better understand the mechanisms of action by which a tumor can form. The ten pathways with respective proteins and genetic alterations are listed in Table 2.

Table 2: Summary of Pathways and Human Oncogenes & Tumor Suppressor Genes

PATHWAY	PROTEIN	ONCOGENE/ TUMOR SUPPRESSORS	GENETIC ALTERATIONS	CONFERRED CAPABILITIES
TGFβ	Myc	Oncogenes	Point mutation, Amplification	Self-sufficiency in growth signals
	BMPR	Tumor Suppressors	Point Mutation	Insensitivity to anti-growth signals
	Smad 2/3	Tumor Suppressors	Point Mutation	Insensitivity to anti-growth signals
	Smad 4	Tumor Suppressors	Point Mutation	Insensitivity to anti-growth signals
	TGFβ R	Tumor Suppressors	Point Mutation	Insensitivity to anti-growth signals
Wnt	B-catenin	Oncogene	Point Mutation	Self-sufficiency in growth signals
	RAR	Oncogene	Amplification	Self-sufficiency in growth signals
	SOX	Oncogene	Amplification, Increased expression	Self-sufficiency in growth signals
	Wnt1	Oncogene	Increased expression	Self-sufficiency in growth signals
	APC	Tumor Suppressors	Point Mutation	Self-sufficiency in growth signals
	Axin	Tumor Suppressors	Point Mutation	Self-sufficiency in growth signals
	α-catenin	Tumor Suppressors	Point Mutation	Tissue invasion & metastasis

	E-cadherin	Tumor Suppressors	Point Mutation	Self-sufficiency in growth signals, Insensitivity to anti-growth signals, Tissue invasion & metastasis
	Wnt5A	Tumor Suppressors	Point Mutation	Self-sufficiency in growth signals
GPCR	Gα	Oncogene	Point Mutation	Self-sufficiency in growth signals
	GPCR	Oncogene	Point Mutation	Self-sufficiency in growth signals
Ras	β-Raf	Oncogene	Point Mutation	Self-sufficiency in growth signals
	Fos/Jun	Oncogene	Increased expression	Evading apoptosis, Self-sufficiency in growth signals
	ILK	Oncogene	Increased expression	Self-sufficiency in growth signals, Tissue invasion & metastasis
	Ras	Oncogene	Point Mutation	Self-sufficiency in growth signals
	RTKs	Oncogene	Point Mutation, Translocation, Amplification, Increased expression	Evading apoptosis, Self-sufficiency in growth signals, Tissue invasion & metastasis, Sustained angiogenesis
	Integrin	Tumor Suppressors	Deletion	Tissue invasion & metastasis
	NF1	Tumor Suppressors	Point Mutation, Deletion	Self-sufficiency in growth signals

	VHL	Tumor Suppressors	Point Mutation	Sustained angiogenesis
Akt	Akt	Oncogene	Point Mutation, Amplification, Increase expression	Evading apoptosis
	Bax	Oncogene	Point Mutation	Evading apoptosis
	FKHR/FOXO	Oncogene	Translocation	Evading apoptosis
	JAK	Oncogene	Point Mutation, Translocation	Evading apoptosis, Self-sufficiency in growth signals
	PI3K	Oncogene	Point Mutation	Evading apoptosis
	Bcl-2	Tumor Suppressors	Translocation	Evading apoptosis
	LKB1	Tumor Suppressors	Point Mutation	Self-sufficiency in growth signals
	PTEN	Tumor Suppressors	Point Mutation, Deletion	Evading apoptosis
	TSC1/TSC2	Tumor Suppressors	Point Mutation	Self-sufficiency in growth signals
Death Receptor	Fas	Oncogene	Point Mutation	Evading apoptosis
Notch	Notch	Oncogene	Translocation	Evading apoptosis
Hedgehog	Gli	Oncogene	Amplification, Translocation	Evading apoptosis, Self-sufficiency in growth signals
	Hedgehog	Oncogene	Point Mutation	Evading apoptosis, Self-sufficiency in growth signals
	Smo	Oncogene	Point Mutation	Evading apoptosis, Self-sufficiency in

				growth signals
	Ptch	Tumor Suppressors	Point Mutation	Evading apoptosis, Self-sufficiency in growth signals
	Su(Fu)	Tumor Suppressors	Point Mutation	Evading apoptosis, Self-sufficiency in growth signals
Cell Cycle	Abl	Oncogene	Translocation	Self-sufficiency in growth signals
	CDK2	Oncogene	Amplification, Increased expression	Self-sufficiency in growth signals
	CDK4	Oncogene	Point Mutation	Self-sufficiency in growth signals
	Cyclin D	Oncogene	Amplification, Translocation	Self-sufficiency in growth signals
	Cyclin E	Oncogene	Amplification	Self-sufficiency in growth signals
	HPV-E6	Oncogene	Viral infection	Self-sufficiency in growth signals
	p 15	Tumor Suppressors	Point Mutation	Self-sufficiency in growth signals
	p 16	Tumor Suppressors	Point Mutation	Self-sufficiency in growth signals
	Rb	Tumor Suppressors	Point Mutation	Self-sufficiency in growth signals
DNA Damage	Aurora A	Oncogene	Amplification, Increased expression	Self-sufficiency in growth signals
	HPV-E6	Oncogene	Viral infection	Evading apoptosis
	MDM2	Oncogene	Amplification	Evading apoptosis

	ARF	Tumor Suppressors	Point Mutation	Self-sufficiency in growth signals
	ATM/ATR	Tumor Suppressors	Point Mutation	Insensitivity to anti-growth signals
	BRCA1	Tumor Suppressors	Point Mutation	Self-sufficiency in growth signals, Insensitivity to anti-growth signals
	Chk1	Tumor Suppressors	Point Mutation	Insensitivity to anti-growth signals
	Chk2	Tumor Suppressors	Point Mutation	Insensitivity to anti-growth signals
	DNA-PK	Tumor Suppressors	Point Mutation	Insensitivity to anti-growth signals
	FANCD2	Tumor Suppressors	Point Mutation	Insensitivity to anti-growth signals
	HIPK2	Tumor Suppressors	Point Mutation	Evading apoptosis, Self-sufficiency in growth signals
	NBS1	Tumor Suppressors	Point Mutation, Deletion	Insensitivity to anti-growth signals
	p 53	Tumor Suppressors	Point Mutation, Deletion	Evading apoptosis, Insensitivity to anti-growth signals

Biomarkers in Cancer

Knowledge of the natural history of cancer progression has advanced in the past decade,

but has been limited in part by the technology available to detect it. Multiple, sequential, parallel, and interconnected changes in cellular machinery over-ride normal biological regulation, and lead to cells becoming neoplastic and invasive. If a cancer could be detected at the incipient stage and its advance halted, we would be able to reduce the mortality associated with the disease[19].

Biomarkers are cellular indicators of the physiological state and also of change during a disease process. Active genes, their respective products, and other organic chemicals made by the cell are unique identifiers that make up the 'molecular signature' of a cell. It has been a challenge to detect these changes in signatures during the early stages of transformation. The signatures may reflect genotoxicity, hyperproliferation, altered patterns of gene expression, hyperplasia, inflammation, aberrant crypt foci, and enzymatic changes that are responses to inherited and environmental causes of cancer[19].

Novel technologies increase our ability to investigate molecular mechanisms of carcinogenesis and might enable us to overcome the challenges in early detection. In this review, we focus on promising molecular signatures, the technologies being developed to detect them, and issues concerning their usefulness[19].

Identification and detection of cancer by pathological techniques are possible only on microscopic examination of the tumor tissue, long after disease onset. Although these techniques are important for prediction of tumor behavior and prognosis, additional methods are necessary for early detection. The usefulness of a biomarker lies in its ability

to provide early indication of disease or the progression of the disease. Biomarkers should be easy to detect, measurable across populations, and amenable to use in one or more of the following settings: detection at an early stage; identification of high-risk individuals; early detection of recurrence; or as intermediate endpoints in chemoprevention[19].

Biomarkers of risk can help identify individuals who are at increased risk of developing cancer, before the biological onset of the disease. These biomarkers are based mainly on inherited or somatically acquired susceptibilities, in the form of altered genes such as MSH2 and MLH in hereditary non-polyposis colorectal cancer, PRB in hereditary retinoblastoma, and BRCA1 and BRCA2 mutated genes that predispose to breast cancer. In these cases, there is an inherent familial predisposition to the development of some type of cancer, although many individuals inheriting mutated genes will not develop cancer. This suggests the involvement of other factors, such as the environment, which could interact with specific genes to initiate cancer. However, risk markers are important in monitoring of individuals and allow early intervention in those who do develop cancer[19].

Genetic inheritance accounts for only a small percentage of cancer incidence in the general population. Biomarkers can detect the outcomes of interaction between genetic susceptibility and the environment and are therefore extremely important for early detection. Theoretically, they could provide the opportunity to intervene during the natural progression of the cancer, to cause inhibition, regression, or even elimination of

the disease. After biological onset, the disease progresses through a preclinical phase before symptoms develop; changes in biomarkers during this phase could be very helpful in early detection, and research has focused on the discovery of biomarkers for this stage[19].

Successful Cancer Products

Oncology products have historically been successful in gaining approval through the FDA, mainly due to the large unmet need. The table (Table 3) below shows all the oncology products approved by the FDA from 2000 to present.

Table 3: FDA Approved Drugs for Oncology[20]

BRAND NAME	PRODUCT NAME	COMPANY	INDICATION	APPROVAL DATE
Abstral	Fentanyl sublingual tablets	ProStraken	Breakthrough cancer pain in opioid-tolerant patients	January 2011
Afinitor	Everolimus	Novartis	Advanced pancreatic neuroendocrine tumors	May 2011
Lazanda	Fentanyl citrate	Archimedes	Management of breakthrough cancer pain	June 2011
Sutent	Sunitinib malate	Pfizer	Pancreatic neuroendocrine tumors	May 2011
Sylatron	Peginterferon alfa-2b	Merck	Melanoma	April 2011

Vandetanib	Vandetanib	Astra Zeneca	Thyroid cancer	April 2011
Yervoy	Ipilimumab	Bristol-Myers Squibb	Metastatic melanoma	March 2011
Zytiga	Abiraterone acetate	Centocor Ortho Biotech	Prostate cancer	May 2011
Halayen	Eribulin mesylate	Eisai	Metastatic breast cancer	November 2010
Herceptin	Trastuzumab	Genentech	Gastric cancer	October 2010
Jevtana	Cabazitaxel	Sanofi Aventis	Prostate cancer	June 2010
Provenge	Sipuleucel-T	Dendreon	Hormone refractory prostate cancer	May 2010
Xgeva	Denosumab	Amgen	Prevention of skeletal-related events in patients with bone metastases from solid tumors	November 2010
Zuplenz	Ondansetron oral soluble film	Strativa Pharmaceuticals	Prevention of post-operative, chemotherapy and radiotherapy induced nausea and vomiting	July 2010

Afinitor	Everolimus	Novartis	Renal cell carcinoma	March 2009
Arzerra	Ofatumumab	GlaxoSmithKline	Chronic lymphocytic leukemia	October 2009
Avastin	Bevacizumab	Genentech	Renal cell carcinoma	July 2009
Cervarix	Human Papillomavirus Bivalent (Types 16 & 18) Vaccine, Recombinant	GlaxoSmithKline	Prevention of cervical cancer and cervical intraepithelial neoplasia caused by HPV types 16 and 18	October 2009
Elitek	Rasburicase	Sanofi Aventis	Management of plasma uric acid levels in adults with malignancies	October 2009
Folotyn	Pralatrexate injection	Allos Therapeutics	Peripheral T-cell lymphoma	September 2009
Istodax	Romidepsin	Gloucester Pharmaceuticals	Cutaneous T-cell lymphoma	November 2009
Onsolis	Fentanyl buccal	BioDelivery Sciences	Management of breakthrough cancer pain	July 2009
Vortrient	Pazopanib	GlaxoSmithKline	Renal cell carcinoma	October 2009
Degarelix	Degarelix for injection	Ferring Pharmaceuticals	Prostate cancer	December 2008
Fusilev	Levoleucovorin	Spectrum Pharmaceuticals	Rescue after high-dose methotrexate therapy in osteosarcoma and to reduce the toxicity of methotrexate	March 2008

Mozobil	Plerixafor injection	Genzyme	Non-Hodgkin's lymphoma and multiple myeloma	December 2008
Sancuso	Granisetron	ProStrakan	Chemotherapy-induced nausea and vomiting	September 2008
Treanda	Bendamustine hydrochloride	Cephalon	Chronic lymphocytic leukemia and B-cell non-Hodgkin's lymphoma	October 2008
Evista	Raloxifene hydrochloride	Eli Lilly	Treatment/prevention of osteoporosis and reduction of breast cancer risk in postmenopausal women	September 2007
Hycamtin	Topotecan hydrochloride	GlaxoSmithKline	Small cell lung cancer	October 2007
Ixempra	Ixabepilone	Bristol-Myers Squibb	Breast cancer	October 2007
Tasigna	Nilotinib hydrochloride monohydrate	Novartis	Chronic myelogenous leukemia	October 2007
Torisel	Temsirolimus	Wyeth	Renal cell carcinoma	May 2007
Tykerb	Lapatinib	GlaxoSmithKline	Breast cancer	March 2007
Gardasil	Quadrivalent human papillomavirus (types 6, 11, 16, 18) recombinant vaccine	Merck	Prevention of cervical cancer associated with human papillomavirus	June 2006
Sprycel	Dasatinib	Bristol-Myers Squibb	Imatinib-resistant chronic myeloid	June 2006

			leukemia	
Sutent	Sunitinib	Pfizer	Kidney cancer and gastrointestinal stromal tumors	January 2006
Vectibex	Panitumumab	Amgen	Colorectal cancer	September 2006
Arranon	Nelarabine	GlaxoSmithKline	T-cell acute lymphoblastic leukemia and T-cell lymphoblastic lymphoma	October 2005
Nexavar	Sorafenib	Bayer/Onyx	Renal Cell Carcinoma	December 2005
Alimta	Pemetrexed for injection	Eli Lilly	Malignant pleural mesothelioma	February 2004
Avastin	Bevacizumab	Genentech	Metastatic carcinoma of the colon or rectum	February 2004
Clolar	Clofarabine	Genzyme	Acute lymphoblastic leukemia in pediatric patients	December 2004
Erbitux	Ceuximab	Imclone	EGFR-expressing, metastatic colorectal cancer	February 2004
Sensipar	Cinacalcet	Amgen	Secondary hyperparathyroidism and hypercalcemia in parathyroid carcinoma patients	March 2004
Tarceva	Erlotinib	Genentech/OSI Pharmaceuticals	Advanced refractory metastatic non-small	November 2004

			cell lung cancer	
Aloxi	Palonosetron	MGI Pharma/Helsinn Healthcare	Prevention of nausea and vomiting associated with emetogenic cancer chemotherapy	August 2003
Bexxar		Corixa	CD20 positive, follicular, non-Hodgkin's lymphoma following chemotherapy relapse	June 2003
Emend	Aprepitant	Merck	Treatment of nausea and vomiting associated with chemotherapy	March 2003
Iressa	Gefitinib	Astra Zeneca	Second-line treatment of non-small-cell lung cancer	May 2003
Plenaxis	Abarelix for injectable suspension	Praecis Pharmaceuticals	Advanced prostate cancer	December 2003
Premarin	Conjugated estrogens	Wyeth	Prevention of postmenopausal osteoporosis and treatment of vasomotor menopause symptoms,	July 2003
UroXatral	Alfuzosin HCl extended-release tablets	Sanofi Aventis	Signs and symptoms of benign prostatic hyperplasia	June 2003
Velcade	Bortezomib	Millennium	Multiple myeloma	May 2003

		Pharmaceuticals	patients who have received at least two prior therapies	
Eloxatin	Oxaliplatin/5-fluorouracil/leucovorin)	Sanofi Aventis	Colon or rectum carcinomas	August 2002
Eligard	Leuprolide acetate	Atrix Laboratories	Palliative treatment of advanced prostate cancer	January 2002
Faslodex	Fulyestrant	Astra Zeneca	Hormone receptor positive metastatic breast cancer	April 2002
Gleevec	Imatinib mesylate	Novartis	Gastrointestinal stromal tumors (GISTs)	February 2002
Neulasta		Amgen	Treatment to decrease the chance of infection by febrile neutropenia in patients receiving chemotherapy	January 2002
SecreFlo	Secretin	Repligen	To aid in the diagnosis of pancreatic dysfunction and gastrinoma	April 2002
Zevalin	Ibritumomab tiuxetan	Biogen Idec	Non-Hodgkin's lymphoma	February 2002
Zometa	Zoledronic acid	Novartis	Multiple myeloma and bone metastases from solid tumors	February 2002
Campath		Berlex	B-cell chronic	May 2001

		Laboratories	lymphocytic leukemia	
Femara	Letrozole	Novartis	First-line treatment of postmenopausal women with locally advanced or metastatic breast cancer	January 2001
Gleevec	Imatinib mesylate	Novartis	Chronic myeloid leukemia	May 2001
Kytril	Granisetron	Roche	Prevention of nausea and vomiting associated with cancer therapy	June 2001
Trelstar LA	Triptorelin pamoate	Debiopharm	Advanced stage prostate cancer	June 2001
Xeloda		Roche	Metastatic colorectal cancer	May 2001
Zometa	Zoledronic acid	Novartis	Hypercalcemia of malignancy	August 2001
Mylotarg	Gemtuzumab ozogamicin	Wyeth	CD33 positive acute myeloid leukemia	May 2000
Trelstar Depot	Triptorelin pamoate	Debiopharm/ Target Research Associates	Palliative treatment of advanced prostate cancer	June 2000
Trisenox	Arsenic trioxide	Cell Therapeutics	Induction of remission and consolidation in patients with acute promyelocytic leukemia	September 2000

Viadur	Leuprolide acetate implant	Alza	Pain relief in men with advanced prostate cancer	March 2000

Even with the approval of these products, there is still a large unmet need in oncology, and thus, biopharmaceutical companies continue to develop potential new cancer treatments.

Chapter 3: Formulating the thesis

The nature of the biopharmaceutical industry introduces many of the challenges discussed above. Each of these challenges incorporates a measure of risk into drug development. Additionally, innovation in technologies for target identification and product screening, an increased number of investigational products in pipelines of biopharmaceutical companies, limited resources for development of new products, increasing costs in the development of new products and increasing uncertainty (risk) has led to a need for portfolio management within medium to large biopharmaceutical companies. Managing the new product pipeline is a series of trade- offs among maximizing expected economic returns, minimizing risk, and maintaining diversity in the product mix for a given level of renewable and nonrenewable corporate resources[21].

A look at the history of portfolio management within the biopharmaceutical industry has revealed many attempts at solving the problem of portfolio management by trying to increase efficiency and improve results, but to date it has not been successful. Many different methods have been used: decision trees, real option valuation, Monte Carlo, and discounted cash flow (DCF). An understanding of the interdependencies within the industry might help to better understand the risks[21]. Four groups of interdependencies within the biopharmaceutical industry have been identified:

Resource dependencies – Example: learning curves lead to decreased development times for similar drugs types

Manufacturing costs dependencies – Example: use of the same facilities for 2 products

Financial return dependencies – Example: cannibalization or complementation

Technical success dependencies – Example: technical success or failure of a drug candidate affects the probability of technical success of an as-yet-untested trailing drug candidate[21]

The level of understanding of technical success interdependencies has not been fully investigated. These interdependencies (correlations) could lead to an overall greater risk to the company's portfolio than previously expected.

Recently, the financial industry has begun to define the systemic risk inherent in hedge funds and markets. This leads one to ask the question: can one use methods of portfolio management and evaluation of risks within finance and apply to the biopharmaceutical industry?

Correlated risk is risk that entails high correlation of markets as well as the degree to which different markets are interconnected. The higher the correlation there is in different markets the greater chance there is for systemic market failure. Since diversifying your investment portfolio with different stocks cannot save you from systemic market failure you have to know how to hedge your bets. Essentially, markets that are negatively

correlated and are fundamentally counteracting each other are the best markets to balance an entire investment portfolio. Does the same methodology apply to biopharmaceutical portfolios? Is it possible to identify the risks that lead to success or failure in drug development, and is it possible to understand the relationships between those risks that might underlie correlations between products?

A better understanding of the risks that lead to success or failure in drug development might encourage more investment in the life sciences and specifically in the biopharmaceutical industry, and a greater awareness of the correlations between risks and products might lead to more informed decision making on a biopharmaceutical portfolio leading increased productivity.

Chapter 4: Methodology

Collecting the data

A dataset was collected from Thomson Reuters. The dataset is the oncology portfolio from a biopharmaceutical company, Genentech Inc. The information collected through this data source were the product's:

- Highest development status (i.e. discovery through launched)
- Therapeutic area
- Mechanism of action
- Target
- Type of technology (e.g. small molecule or biologic)

For purposes of this analysis, all products still in discovery were excluded because the risk factors are very different than clinical stage products and there are fewer publications on those risks.

The remainder of the data was collected through literature research. Potential and known risks factors were identified. These risk factors were narrowed down to the following:

- Expected return
- Established mechanisms of action
- Proven animal model

- Available biomarker for disease (identified at FDA approval)
- Available biomarker for dosing (identified at FDA approval)
- Companion diagnostic developed (approval with product)
- Established endpoints for primary indication
- FDA prior experience with primary indication
- FDA prior experience with mechanism of action (pathway)
- Reimbursement potential
- Established standard of care for primary indication
- Presence of dose-response relationship

The assumptions for each of these risk factors are as follows:

- Therapeutic area: for those products without a primary indication specified, a generic term "cancer" was used. For purposes of estimating the market size for "cancer," an average incidence per type of cancer was used. This data was collected from the American Cancer Society. For those products with "solid tumor" specified for therapeutic area, the same approach was used for estimating market size.
- Expected return: For launched products, cumulative sales to date were collected from company financial statements and rounded to the nearest million. For products still in development, the following assumptions were build into the forecasted expected return: 7 years of market exclusivity, 30% market penetration, 3 years to peak sales, 10% discount rate and pricing based on an

average of oral oncology products (source: pharmacytimes.com, "Oral Oncology Therapies: Specialty Pharmacy's Newest Challenge). The market size used for those products with the generic "cancer" was an average of all cancer cases taken from the American Cancer Society 2011 statistics.

- Established mechanisms of action: if information could not be found through a literature search, then it was assumed that a mechanism of action was not established.
- Proven animal model: if a mechanism of action is not established, then it was assumed that an animal model was not proven. Additionally, if information could not be found through a literature search, then it was assumed that an animal model was not proven.
- Available biomarker for disease (identified at FDA approval): if information could not be found through a literature search, then it was assumed that a disease biomarker was not available.
- Available biomarker for dosing (identified at FDA approval): if information could not be found through a literature search, then it was assumed that a dosing biomarker was not available.
- Companion diagnostic developed (approved with product): if information could not be found through either clinicaltrials.gov or company website, then it was assumed that a companion diagnostic was not in co-development.
- Established endpoints for primary indication: endpoints were considered established if an FDA guidance exists

- FDA prior experience with primary indication: it was assumed that the FDA had prior experience with a disease area if a guidance was provided or if another drug product has been approved.
- FDA prior experience with mechanism of action (pathway): it was assumed that the FDA had prior experience with a biological pathway if a guidance was provided or if another drug product has been approved with the same mechanism of action.
- Reimbursement potential: it was assumed that all products would be reimbursed due to the unmet need in oncology therapies.
- Established standard of care for primary indication: if information could not be found through a literature search, then it was assumed that a standard of care was not established.
- Presence of dose-response relationship: if information could not be found through a literature search, then it was assumed that a dose-response relationship was not present.

Analysis

The first step of the analysis is to observe the data for any variables that resulted in the same value for all outcomes, success and failure, and to remove those variables from the dataset. For example, if all products, regardless of success or failure, in the dataset had the same value for an independent variable, that independent variable should be removed

from the analysis. This is done because those variables will not contribute to the regression model or the overall outcome.

The second step of the analysis is to count the raw number of outcomes for each independent variable by the overall outcome of success or failure. Then, by normalizing these raw values and calculating the delta between those normalized percentages for successes and failures, one can eliminate other independent variables that will not have large contributions to regression model. The idea in eliminating independent variables that will not largely contribute is to ultimately build a regression model that only contains variables that are significant to the outcome.

Logistic regression is part of a category of statistical models called generalized linear models. Logistic regression allows one to predict a discrete outcome from a set of variables that may be continuous, discrete, dichotomous, or a mix of any of these. Generally, the dependent variable is dichotomous, such as success/failure. In instances where the independent variables are a categorical, or a mix of continuous and categorical, logistic regression is used.

The dependent variable in logistic regression is usually dichotomous; that is, the dependent variable can take the value 1 with a probability of success q, or the value 0 with probability of failure 1-q. This type of variable is called a Bernoulli (or binary) variable. As mentioned previously, the independent or predictor variables in logistic regression can take any form. That is, logistic regression makes no assumption about the

distribution of the independent variables. They do not have to be normally distributed, linearly related or of equal variance within each group. The relationship between the predictor and response variables is not a linear function in logistic regression; instead, the logistic regression function is used, which is the logit transformation of q:

$$\theta = \frac{e^{(\alpha+\beta_1 x_1+\beta_2 x_2+\ldots+\beta_i x_i)}}{1+e^{(\alpha+\beta_1 x_1+\beta_2 x_2+\ldots+\beta_i x_i)}}$$

Where *a* is equal to the constant of the equation and *b* is equal to the coefficient of the predictor variables. The goal of logistic regression is to correctly predict the category of outcome for individual cases using the tightest model. To accomplish this goal, a model is created that includes all predictor variables that are useful in predicting the response variable. Several different options are available during model creation. Variables can be entered into the model in the order specified by the researcher or logistic regression can test the fit of the model after each coefficient is added or deleted, called stepwise regression.

Stepwise regression is used in the exploratory phase of research but it is not recommended for theory testing (Menard 1995). Theory testing is the testing of a-priori theories or hypotheses of the relationships between variables. Exploratory testing makes no a-priori assumptions regarding the relationships between the variables; thus the goal is to discover relationships.

Backward stepwise regression appears to be the preferred method of exploratory analyses, where the analysis begins with a full or saturated model and variables are eliminated from the model in an iterative process. The fit of the model is tested after the elimination of each variable to ensure that the model still adequately fits the data. When no more variables can be eliminated from the model, the analysis has been completed. For this analysis, backward stepwise regression was used.

There are two main uses of logistic regression. The first is the prediction of group membership. Since logistic regression calculates the probability or success over the probability of failure, the results of the analysis are in the form of an odds ratio. For example, logistic regression is often used in epidemiological studies where the result of the analysis is the probability of developing cancer after controlling for other associated risks. Logistic regression also provides knowledge of the relationships and strengths among the variables (e.g., smoking 10 packs a day puts you at a higher risk for developing cancer than working in an asbestos mine).

The process by which coefficients are tested for significance for inclusion or elimination from the model involves several different techniques. The first is the Wald test. A Wald test is used to test the statistical significance of each coefficient (b) in the model. A Wald test calculates a *Z* statistic, which is:

$$Z = \frac{\hat{B}}{SE}$$

This z value is then squared, yielding a Wald statistic with a chi-square distribution. However, several authors have identified problems with the use of the Wald statistic. Menard (1995) warns that for large coefficients, standard error is inflated, lowering the Wald statistic (chi-square) value. Agresti (1996) states that the likelihood-ratio test is more reliable for small sample sizes than the Wald test. The second method for testing coefficients for significance is the likelihood-ratio test. The likelihood-ratio test uses the ratio of the maximized value of the likelihood function for the full model (L_1) over the maximized value of the likelihood function for the simpler model (L_0). The likelihood-ratio test statistic equals:

$$-2\log(\frac{L_0}{L_1}) = -2[\log(L_0) - \log(L_1)] = -2(L_0 - L_1)$$

This log transformation of the likelihood functions yields a chi-squared statistic. This is the recommended test statistic to use when building a model through backward stepwise elimination. The third method for testing coefficients for significance is the Hosmer-Lemshow goodness of tit test. The Hosmer-Lemshow statistic evaluates the goodness-of-fit by creating 10 ordered groups of subjects and then compares the number actually in the each group (observed) to the number predicted by the logistic regression model (predicted). Thus, the test statistic is a chi-square statistic with a desirable outcome of non-significance, indicating that the model prediction does not significantly differ from the observed. The 10 ordered groups are created based on their estimated probability;

those with estimated probability below 0.1 form one group, and so on, up to those with probability 0.9 to 1.0. Each of these categories is further divided into two groups based on the actual observed outcome variable (success, failure). The expected frequencies for each of the cells are obtained from the model. If the model is good, then most of the subjects with success are classified in the higher deciles of risk and those with failure in the lower deciles of risk[22].

StatPlus® software tool was used to perform the logistic regression on Excel.

Chapter 5: Results

A total of forty-five data points were available from the Thomas Reuters Genentech Inc. oncology dataset. After removing all discovery stage and clinical stage products, sixteen products were left for the analysis that had either launched or been discontinued. All raw data can be found in Appendix A.

During the first step of the analysis, three independent variables were removed because the values were the same for all data points. The variables removed were biomarker for dose, reimbursement potential, and presence of standard of care. These variables would not contribute to the regression model or the outcome.

The results of the raw number and normalization analysis led to the elimination of four more independent variables. These variables were available biomarker for disease (identified at FDA approval), established endpoints for primary indication, FDA prior experience with primary indication and type of technology (biologic vs. small molecule).

The threshold for elimination was a delta of less than15%. All variables with a delta of less than 15% were eliminated from final regression model. All variables with a delta of greater than or equal to 15% were included in the final regression model. The output of this analysis is shown in Table 4.

Table 4: Raw Number, Normalization and Delta Calculation for Independent Variables

	Established Mechanism of Action	Animal Model	Biomarker for disease	Companion Diagnostic	Established Endpoints	FDA prior experience-disease	FDA prior experience-pathway	Type of technology	Dose-response relationship
Raw Number									
Launched	5	5	2	1	5	5	5	3	2
Failed	8	9	5	4	10	10	6	8	1
Normalized									
Launched	100%	100%	40%	20%	100%	100%	100%	60%	40%
Failed	73%	82%	45%	36%	91%	91%	55%	73%	9%
Delta	27%	18%	-5%	16%	9%	9%	45%	-13%	31%

The logistic regression was performed with the remaining five variables (established mechanism of action, presence of animal model, companion diagnostic, FDA prior experience with pathway and dose-response relationship). Output from the logistic regression can be found in Appendix B.

There were 11 observations of the dependent variable equal to 0, failure, and there were 5 observations of the dependent variable equal to 1, launched.

The chi-square value was 7.738 with the degrees of freedom equal to 5 and with a p-value of 0.17.

The regression coefficients for each independent variable were as follows:

- Mechanism of action: $\beta = 17.10$ with p-value of 0.99
- Animal model: $\beta = -16.65$ with p-value of 0.99

- Companion diagnostic: β = -1.79 with p-value of 0.29
- FDA prior experience with pathway: β = 16.65 with p-value of 0.99
- Dose-response relationship: β = 1.94E-14 with p-value of 1.00

Since all p-values are much greater than 0.05, none of these variables directly contribute to the outcome of launched or failure.

The odds ratio for each independent variable were as follows:

- Mechanism of action: 26,616,199
- Animal model: 0
- Companion diagnostic: 0.1667
- FDA prior experience with pathway: 17,066,840
- Dose-response relationship: 1

These results suggest that the variables mechanism of action and FDA prior experience with pathway could increase the odds of the outcome (i.e. odds ratio is greater than 1) when the value of the independent variable is increased by 1 unit.

Chapter 6: Discussion

This analysis was limited by data available for the chosen company. A larger dataset incorporating multiple companies might have provided a more robust analysis and would have removed any confounding effects directly linked to the company. Logistic regression tends to systematically overestimate odds ratios or beta coefficients when the sample size is less than about 500. With increasing sample size, the magnitude of overestimation diminishes and the estimated odds ratio asymptotically approaches the true population value. In a single study, overestimation due to small sample size might not have any relevance for the interpretation of the results, since it is much lower than the standard error of the estimate. However, if a number of small studies with systematically overestimated effects are pooled together without consideration of this effect, an effect may be perceived when in reality it does not exist. A minimum of 10 events per independent variable has been recommended[4].

The chi-square value was 7.738 with a p-value of 0.17. The most common assessment of overall model fit in logistic regression is the goodness-of-fit test (G), which is simply the chi-square difference between the null model (i.e. with the constant only) and the model containing the one or more predictors. This is one use of the likelihood ratio test between two nested models (referred to as "chi-square" in StatPlus binary logistic regression output). It is an assessment of the improvement of the fit between the predicted and the observed values on Y by adding the predictors to the model.

Each of the regression coefficients describes the size of the contribution of that risk factor. A positive regression coefficient means that the explanatory variable increases the probability of the outcome, while a negative regression coefficient means that the variable decreases the probability of that outcome; a large regression coefficient means that the risk factor strongly influences the probability of that outcome, while a near-zero regression coefficient means that that risk factor has little influence on the probability of that outcome.

One difficultly with a larger dataset is collecting all of the risk variable information for each product. Either some the data might not be available or there might be conflicting evidence that contradicts each other. In both cases, these data points would need to be excluded from the overall analysis because of the missing inputs.

Other risk variables could have been included in the model. For example, other regulatory, manufacturing and commercial risks could have been included in the analysis. These variables might have provided additional insights into predicting success or failure of products. However, the more variables included, the greater likelihood that some of those variables are not needed. That is, the variables might not have an effect or they might be related such that they confound the results.

These types of analyses help to identify the risk factors that play the largest role in the outcome of a product in drug development. As more evidence become available, the data might be able to assist in biopharmaceutical portfolio decision making within

biopharmaceutical companies. A major concern with biopharmaceutical companies is their decreasing levels of productivity and high attrition rates. As we start to better understand the risks that are involved in drug development and to determine which risks contribute to the success or failure of a product, then companies can make more informed decisions in portfolio management. This should lead to companies funding products that have a higher probability of success of gaining approval from the FDA and being successful in the market.

Other types of analyses that would augment this logistic regression analysis include portfolio optimization with specific constraints around the risk factors and qualitative research with interviews from biopharmaceutical executives. The combination of all of the analyses could lead to further insights into with risk factors might predict the success or failure of a biopharmaceutical product.

Additionally, this data, especially combined with other similar analyses, should educate investors on the risks involved with investing in the life sciences. As investors become available of the risk and start to understand the risk better, they should be more motivated to invest. This assumes that a reasonable return is provided.

Bibliography

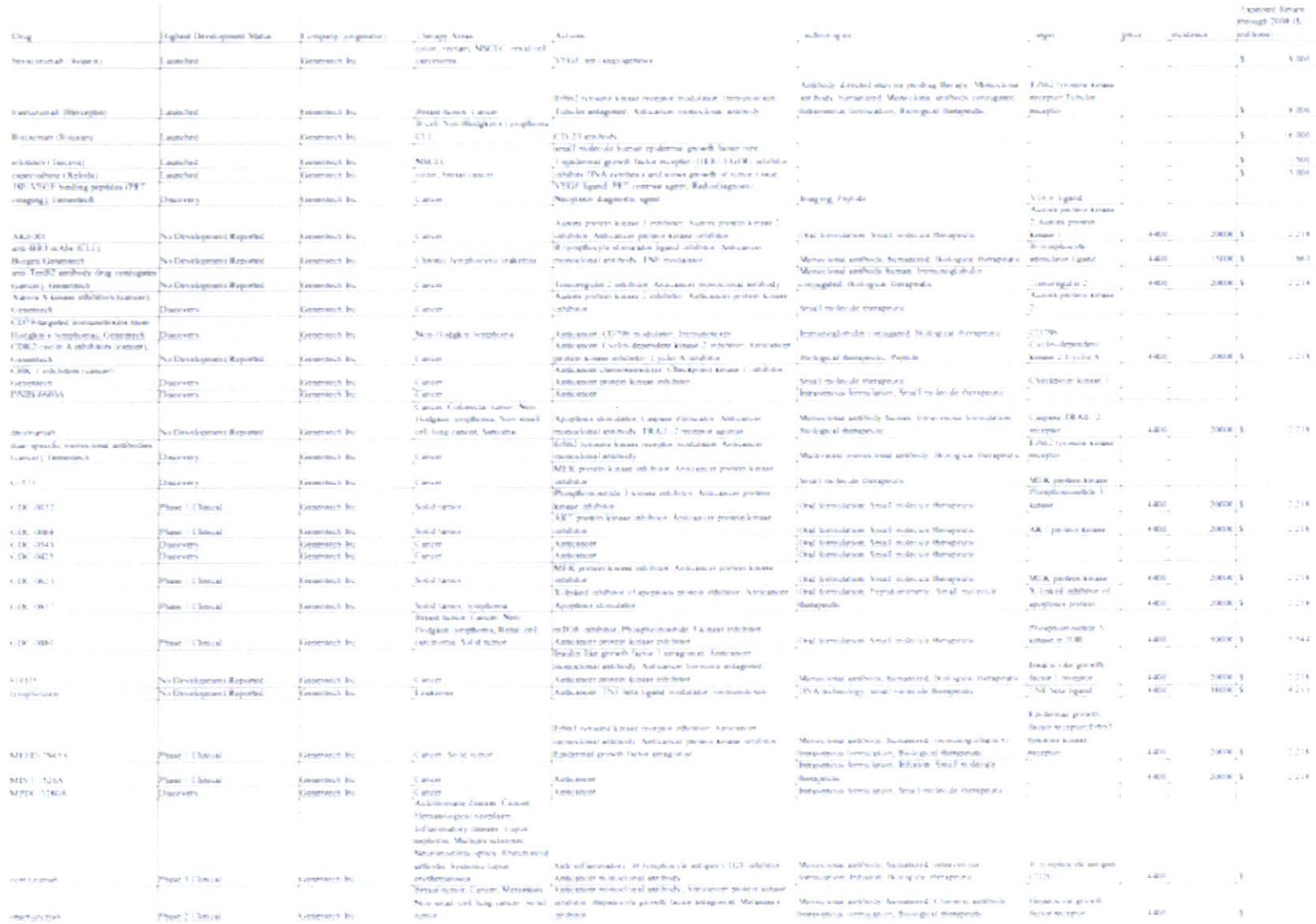
1. BCG. Capitalizing on the Crisis: New Ways to Create Value in Biopharma.
2. www.nih.gov
3. Lee D and Dibner M. The rise of venture capital and biotechnology in the US and Europe. Nature. 2005.
4. www.wikipedia.com
5. The Kauffman Foundation
6. Center for Venture Research
7. Angel Capital Education Foundation
8. www.investopedia.com
9. Stewart J, et al. Nature Biotechnology 2001, Volume 19.
10. Bloomberg interview, What Venture Capital Wants out of Health Care: Q&A with Mike Carusi of Advanced Technology Ventures, 04Nov99.
11. Lee D and Dibner M. The rise of venture capital and biotechnology in the US and Europe. Nature. 2005.
12. Cambridge Associates
13. Dickson, Michael. The Cost of New Drug Discovery and Development. Discovery Medicine. 2000.
14. Lavole, The Source of Comparative Advantage in the Biotechnology Industry: A Real Options Approach, AAEA Annual Meeting, 1999.
15. www.fda.gov
16. Source for phases of drug development
17. Business Insights. Winning R&D Productivity Strategies: Exploiting innovation, licensing and outsourcing opportunities.
18. Sculpher M, Claxton K. Establishing the cost-effectiveness of new pharmaceuticals under conditions of uncertainty--when is there sufficient evidence? Value Health. 2005 Jul-Aug;8(4): 433-46.
19. Srinivas P. et al. The Lancet Oncology, November 2001; Volume 2, Issue 11, Pages 698 – 704.

20. www.centerwatch.com

21. Gary E. Blau, Joseph F. Pekny, Vishal A. Varma, and Paul R. Bunch. Managing a Portfolio of Interdependent New Product Candidates in the Pharmaceutical Industry. J Prod Innov Manag 2004; 21: 227–245.

22. www.oxfordjournals.org

Appendix A: Raw Data for Genentech Inc.'s Oncology



Drug	Highest Development Status	Company Comparator	Therapy Areas	Actions	Technologies	Target	price	Incidence	Expected Return through 2008 (5 m Euros)	
			Breast cancer, Cancer, Non-small cell lung cancer, Ovary cancer, Pancreas cancer, Prostate cancer							
	Phase I Clinical	Genentech Inc	Stomach cancer							
Genentech	Discovery	Genentech Inc	Cancer							
			Chronic lymphocytic leukemia, Hematological neoplasms, Non-Hodgkin lymphoma							
	No Development Reported	Genentech Inc								
	Discovery	Genentech Inc	Cancer							
	No Development Reported	Genentech Inc	Solid cancer							
	Phase 2 Clinical	Genentech Inc	Cancer, Non-small cell lung cancer, Solid tumors							
	No Development Reported	Genentech Inc	Cancer							
	Phase I Clinical	Genentech Inc	Cancer, Multiple myeloma							
	Phase I Clinical	Genentech Inc	Cancer, Hematological neoplasm							
	Discovery	Genentech Inc	Cancer							
	Phase I Clinical	Genentech Inc	Breast cancer, Cancer							
	No Development Reported	Genentech Inc	Cancer							
	Discovery	Genentech Inc	Cancer							
	Phase 1 Clinical	Genentech Inc	Breast cancer, Cancer, Colorectal cancer, Ovary cancer, Pancreatic ductal adenocarcinoma							

Drug	Established MOA	MOA Dummy	Phase [illegible] (model)	Animal Dummy	Available biomarker for target (at approval)	Biomarker Target Dummy	Available biomarker for [illegible]	Biomarker [illegible] Dummy	[illegible] (at approval)	[illegible] Dummy	[illegible]	[illegible] Dummy	FDA [illegible] with Genentech	FDA [illegible] Dummy
bevacizumab (Avastin)	yes		yes		no		no		no		yes		yes	
trastuzumab (Herceptin)	yes		yes		yes		no		yes		yes		yes	
Rituximab (Rituxan)	yes		yes		no		no		no		yes		yes	
erlotinib (Tarceva)	yes		yes		no		no		no		yes		yes	
capecitabine (Xeloda)	yes		yes		yes		no		no		yes		yes	
[illegible] VEGF binding peptides (PET imaging) (Genentech)														
ARQ 197	no		no		no		no		no		yes		yes	
anti-HER3 mAb (CE1) Biogen Genentech	yes		yes		yes		no		yes		yes		yes	
anti-TenB2 antibody drug conjugates (cancer) (Genentech)	no		no		no		no		no		no		no	
Aurora A kinase inhibitors (cancer) (Genentech)														
CD20 targeted [illegible] for non-Hodgkin's lymphoma (Genentech)														
CDK/cyclin [illegible] inhibitors (cancer) (Genentech)	yes		yes		yes		no		yes		yes		yes	
CHK[illegible] inhibitors (cancer) (Genentech)														
[illegible]														
[illegible]mab	yes		yes		no		no		no		yes		yes	
dual specific monoclonal antibodies (cancer) (Genentech)														
[illegible]														
GDC-[illegible]	yes		yes		yes		yes		yes		yes		yes	
GDC-[illegible]	yes		yes		yes		no		yes		yes		yes	
GDC-[illegible]														
GDC-[illegible]														
GDC-[illegible]	yes		yes		yes		no		yes		yes		yes	
GDC-[illegible]	yes		yes		no		no		no		yes		yes	
GDC-[illegible]	yes		yes		yes		no		yes		yes		yes	
[illegible]	yes		yes		yes		no		yes		yes		yes	
[illegible]	yes		yes		no		no		no		yes		yes	
MEHD7945A			[illegible]		[illegible]						[illegible]		[illegible]	
MINT1526A			[illegible]		[illegible]						[illegible]		[illegible]	
MPDL3280A														
[illegible]mab			[illegible]		[illegible]						[illegible]		[illegible]	
[illegible]mab			[illegible]		[illegible]						[illegible]		[illegible]	

Drug	FDA-approved with indication	PDA pathway Diagnosis	Biomarker-driven approval	Recent Diagnosis	Biologic vs Small Molecule	Bio vs SM Diagnosis	Standard of care established	SOC Diagnosis	Dose response demonstrated	Dose Response Diagnosis
[illegible] (Avastin)	yes	1	yes		biologic		yes		yes	1
[illegible] (Herceptin)	yes	1	yes		biologic		yes		yes	1
[illegible] (Rituxan)	yes	1	yes		biologic		yes		yes	1
erlotinib (Tarceva)	yes	1	yes		small molecule		yes		yes	1
cetuximab (Erbitux)	yes	1	yes		small molecule		yes		yes	1
18F-YE105 binding peptides (PET imaging), Genentech										
ABL001	no	0	yes		small molecule		yes		yes	1
anti-BR3 mAbs (CLL),										
Biogen/Genentech	no	0	yes		biologic		yes		yes	1
anti-Her2 antibody drug conjugates (cancer), Genentech	no	0	yes		biologic		yes		yes	1
Aurora A kinase inhibitors (cancer), Genentech										
CD79 targeted immunotoxins (non-Hodgkin's lymphomas), Genentech										
CHK2 tyrosine kinase inhibitors (cancer), Genentech	yes	1	yes		biologic		yes		yes	1
CHK1 inhibitors (cancer), Genentech										
DYRK inhibitors										
dual specific monoclonal antibodies (cancer), Genentech	no	0	yes		biologic		yes		yes	1
[illegible]										
GDC-0032	yes	1	yes		small molecule		yes		yes	1
GDC-0068 GDC-0349	yes	1	yes		small molecule		yes		yes	1
GDC-0425										
GDC-0623	yes	1	yes		small molecule		yes		yes	1
GDC-0941	yes	1	yes		small molecule		yes		yes	1
GDC-0980	yes	1	yes		small molecule		yes		yes	1
IC1465	yes	1	yes		biologic		yes		yes	1
[illegible]	no	0	yes		small molecule		yes		yes	1
MEDI-7944A		0		0		0		0		0
MEDI-1179A MEDI-1203A		0		0		0		0		0
[illegible]		0		0		0		0		0
[illegible]		0		0		0		0		0

Drug	FDA-approved for this pathway	FDA pathway Targets	Bevacizumab (anti-VEGF)	Bcr-Abl Targets	Erlotinib (anti-EGFR)	Bcr-Abl TKI Targets	Sunitinib (multi-target)	VEGF Targets	Dasatinib (multi-target)	Abl Targets
[illegible]		0			0		0		0	
Panitumumab (anti-EGFR) monoclonal antibody										
PRO-131921 radiolabeled macrocyclic chelating agent	yes	yes			Phase I		yes		yes	
SCC-7347	yes	yes			Phase I		yes		yes	
SCC-7414		0			0		0		0	
SCC-7415	yes	yes			Small molecule		Orphan		yes	
SCC-7494		0			0		0		0	
SCC-7431		0			0		0		0	
Cetuximab/Erbitux (EMD)										
Cetuximab analogue		0			0		0		0	
Matuzumab (EMD-72000)	yes	yes			Phase I		yes		yes	
Cetuximab (IMC-C225/MAb-225) antibody drug conjugate (ADC)										
Chemotherapy		0			0		0		0	

Appendix B: Output from Binary Logistic Regression on StatPlus

Binary Logistic Regression

Summary

VAR	*Mean*	*Standard Deviation*
MOA Dummy (#1)	0.8125	0.39031
Animal Dummy (#2)	0.875	0.33072
Dx Dummy (#3)	0.3125	0.46351
FDA pathway Dummy (#4)	0.6875	0.46351
Dose-Response Dummy (#5)	0.1875	0.39031

Total number of observations Dependent Variable=0	11
Total number of observations Dependent Variable=1	5

Iterations	
Null model -2 Log Likelihood	19.87476
Full model -2 Log Likelihood	13.53057
Full model -2 Log Likelihood	12.58242
Full model -2 Log Likelihood	12.29551
Full model -2 Log Likelihood	12.1946
Full model -2 Log Likelihood	12.15801
Full model -2 Log Likelihood	12.14463
Full model -2 Log Likelihood	12.13971
Full model -2 Log Likelihood	12.1379
Full model -2 Log Likelihood	12.13724
Full model -2 Log Likelihood	12.13699
Full model -2 Log Likelihood	12.1369
Full model -2 Log Likelihood	12.13687
Full model -2 Log Likelihood	12.13686
Full model -2 Log Likelihood	12.13685
Process Converged	

Overall Model Fit

Chi-square	7.73791	*Degrees Of Freedom*	5	*p-level*	0.171

Regression Statistics

VAR	*Beta*	*Standard Error*	*p-level*
MOA Dummy (#1)	17.09703	2,212.64066	0.99383
Animal Dummy (#2)	-16.65265	2,965.22533	0.99552
Dx Dummy (#3)	-1.79176	1.68325	0.28712
FDA pathway Dummy (#4)	16.65265	1,203.70074	0.98896
Dose-Response Dummy (#5)	1.94471E-14	1.73205	1.
Intercept	-16.40388		

Odds Ratios

VAR	*Odds Ratios*	*LCL*	*UCL*
MOA Dummy (#1)	26,616,198.92769	0.E+0	#N/A
Animal Dummy (#2)	0.	0.E+0	#N/A
Dx Dummy (#3)	0.16667	0.00615	4.51503
FDA pathway Dummy (#4)	17,066,840.06818	0.E+0	#N/A
Dose-Response Dummy (#5)	1.	0.03355	29.80927

Mechanism of Action	**Animal Model**	**Companion Dx**	**FDA pathway**	**Dose-Response**	**Dependent Variable**	**p-level**
1.	1.	2.04643E-17	1.	1.	1.	0.666
1.	1.	1.	1.	1.04761E-17	1.	0.
1.	1.	2.04643E-17	1.	1.	1.	0.666
1.	1.	2.04643E-17	1.	1.04761E-17	1.	0.666
1.	1.	2.04643E-17	1.	1.04761E-17	1.	0.666
3.0954E-17	6.37511E-17	2.04643E-17	-2.69966E-17	1.04761E-17	0.E+0	
1.	1.	1.	-2.69966E-17	1.04761E-17	0.E+0	
3.0954E-17	6.37511E-17	2.04643E-17	-2.69966E-17	1.04761E-17	0.E+0	
1.	1.	1.	1.	1.04761E-17	0.E+0	0.
1.	1.	2.04643E-17	-2.69966E-17	1.04761E-17	0.E+0	
1.	1.	1.	1.	1.04761E-17	0.E+0	0.
1.	1.	2.04643E-17	-2.69966E-17	1.04761E-17	0.E+0	
1.	1.	1.	1.	1.04761E-17	0.E+0	0.
1.	1.	2.04643E-17	1.	1.	0.E+0	0.666
3.0954E-17	1.	2.04643E-17	1.	1.04761E-17	0.E+0	
1.	1.	2.04643E-17	1.	1.04761E-17	0.E+0	0.666

	Predicted Y		
observed	0	1	Total
0	9	2	11
1	1	4	5
Total	10	6	16